A novel Parkinson's Disease Diagnosis Index using higher-order spectra features in EEG signals

Article *in* Neural Computing and Applications · August 2018

DOI: 10.1007/s00521-016-2756-z

CITATIONS
36

READS
763

3 authors:

YUVARAJ RAJAMANICKAM
Nanyang Technological University
26 PUBLICATIONS **644** CITATIONS
SEE PROFILE

U Rajendra Acharya
Ngee Ann, Singapore University of Social Science, Singapore; University of Malaya…
723 PUBLICATIONS **30,717** CITATIONS
SEE PROFILE

Yuki Hagiwara
Fraunhofer IKS
59 PUBLICATIONS **3,488** CITATIONS
SEE PROFILE

Neural Comput & Applic
DOI 10.1007/s00521-016-2756-z

ORIGINAL ARTICLE

A novel Parkinson's Disease Diagnosis Index using higher-order spectra features in EEG signals

Rajamanickam Yuvaraj[1,3] · **U. Rajendra Acharya**[2] · **Yuki Hagiwara**[2]

Received: 4 April 2016 / Accepted: 23 November 2016

Abstract Higher-order spectra (HOS) is an efficient feature extraction method used in various biomedical applications such as stages of sleep, epilepsy detection, cardiac abnormalities, and affective computing. The motive of this work was to explore the application of HOS for an automated diagnosis of Parkinson's disease (PD) using electroencephalography (EEG) signals. Resting-state EEG signals collected from 20 PD patients with medication and 20 age-matched normal subjects were used in this study. HOS bispectrum features were extracted from the EEG signals. The obtained features were ranked using *t* value, and highly ranked features were used in order to develop the PD Diagnosis Index (PDDI). The PDDI is a single value, which can discriminate the two classes. Also, the ranked features were fed one by one to the various classifiers, namely decision tree (DT), fuzzy K-nearest neighbor (FKNN), K-nearest neighbor (KNN), naive bayes (NB), probabilistic neural network (PNN), and support vector machine (SVM), to choose the best classifier using minimum number of features. We have obtained an optimum *mean* classification accuracy of 99.62%, *mean* sensitivity and specificity of 100.00 and 99.25%, respectively, using the SVM classifier. The proposed PDDI can aid the clinicians in their diagnosis and help to test the efficacy of drugs.

Keywords Electroencephalogram · Parkinson's disease · Higher-order spectra · Machine learning algorithms · Diagnosis index

1 Overview

At birth, the brain contains as many nerve cells called neurons; the number of neurons at birth is the most that will ever be present in the brain. Unlike other cells of our body, such as skin or bone, the nerve cells cannot self-repair. Therefore, as we get older, the neurons die due to aging and are irreplaceable. Parkinson's disease (PD) is a slowly advancing neurological condition that destroys the dopamine neurons in the substantia nigra pars compacta, which affects communication pathways of the brain. The PD usually affects people aged 50 or older and has affected approximately 7–10 million people worldwide to date, according to the WHO statistics [1]. In the next 25 years, the number of the PD is expected to increase due to the rise in the proportion of elderly people [1]. The cardinal symptoms of PD are tremor, muscle stiffness, bradykinesia (slowness of movement), unstable posture, or diminished balance and coordination and dysphonia (voice disorders). In addition, the PD is also identified by with the occurrence of non-motor symptoms including cognitive dysfunction. However, the diagnosis of PD based on the clinical symptoms remains complicated, when there are no significant motor or non-motor symptoms. On the other hand, the detection of PD from the neuroscience viewpoint provides us with another way to explore the neuronal mechanisms and improve diagnostic findings [2].

✉ Rajamanickam Yuvaraj
yuva2257@gmail.com

[1] Department of Biomedical Engineering, Sri Sivasubramaniya Nadar (SSN) College of Engineering, Rajiv Gandhi Salai, Kalavakkam, Tamilnadu 603110, India

[2] Department of Electronics and Computer Engineering, Ngee Ann Polytechnic, Clementi 599489, Singapore

[3] Present Address: Neural Systems Laboratory, Department of Biomedical Engineering, University of Kentucky, 143 Graham Ave, Lexington, KY 40506-0108, USA

Published online: 09 December 2016

Researchers have proposed various noninvasive methods to detect PD using voice, gait, and wearable sensors [3–5]. Chen et al. [59] used the feature reduction method to exclude the redundant information in the original PD voice signal, which was integrated with fuzzy classifier for PD diagnosis. They have achieved a *mean* classification accuracy of 96.07% [6]. Zuo et al. [7] presented a study based on particle swarm optimization enhanced with fuzzy for the PD detection and obtained the highest accuracy of 97.47%. Ma et al. [4] combined extreme learning machine with subtractive clustering features algorithm to detect PD, which resulted with the best *mean* accuracy of 99.49%. Daliri [5] used gait information and employed feature discriminant ratio based on short-time Fourier transform and achieved maximum classification accuracy of 91.20%. In recent years, detection of PD using electroencephalogram (EEG) signals has drawn significant attention of numerous researches as the PD is a complicated neurodegenerative disease and there could be lot of information underlying neural mechanisms [1] compared to other modalities. It has been extensively shown that the EEG analysis, such as linear or nonlinear methods, could depict more global indices of brain functions, which can reflect the disturbed subcortico-cortical mechanisms in patients with PD. Various researchers have revealed abnormalities in the PD patient's EEG signal by using conventional spectral method (e.g., fast Fourier transform [FFT]), time–frequency analysis (e.g., wavelet), or nonlinear time series methods (e.g., correlation dimension) [1, 8–12]. Pezard et al. (2001) found a significant decrease in EEG spectral power between the PD patients and normal subjects [10]. Muller et al. [13] revealed a significant difference in the resting-state EEG signals in PD patients compared to the normal subjects using correlation dimension. Han et al. [1] found that the PD patient's EEG signals are characterized by higher entropy in the frequency domain. Thus, the understanding of the neural basis in PD is essential, both from a prognostic perspective and for the development of targeted therapeutic strategies. However, it is still unclear which measures can be more useful to reveal more important information regarding brain dysfunctions. Furthermore, the EEG has been increasingly used to recognize the cortical integrative functions and their subcortical driving structures. It has also been proved that computing EEG signals can provide a vital biomarker for many neurological disorders such as epilepsy, schizophrenia, Alzheimer's disease [14–17].

Over the years, several linear methods have been used to analyze the EEG signals and compute the central nervous system activity in PD patients [1, 10, 13]. However, such methods are not effective to identify the subtle variations in EEG signals due to their complex, nonlinear, and chaotic nature. Even though frequency-domain methods are used, the accuracy of spectral information decreases signal-to-noise ratio [18]. Despite that, nonlinear algorithms are widely used to unearth the hidden signatures from EEG signals [19]. Moreover, it has been shown that the development of PD is associated with slowing of EEG, reduction of its complexity, and the presence of the perturbations in EEG synchrony. This essential and hidden information can be evaluated by understanding the nonlinear components present in the EEG signals since it captures the momentary variations related to the properties such as similarity, predictability, reliability, and sensitivity of the signal. In recent years, subtle changes in biosignals have been extracted via the widely used higher-order spectra (HOS) method [20–23]. The HOS-based bispectrum methods have been used for various biomedical applications such as epilepsy diagnosis [20], sleep stages [21], cardiac abnormalities [22], effective computing [23]. However, such studies have not yet been performed on PD patient's EEG signals. This work, therefore, aims to extract the hidden changes in the EEG signals to help the automated classification of PD.

A description of materials (i.e., participants and EEG recording) used in this study is given in Sect. 2. In Sect. 3, the detailed description of methodology (includes preprocessing, HOS features, ranking, and classification) is presented, followed by experimental results in Sect. 4. We have also proposed "PD diagnosis index," to find whether the recorded EEG signal belongs to normal or PD patient using a single numeric value. The interpretation of the results and conclusion of this study are given under Sects. 5 and 6, respectively.

2 Materials

2.1 Participants

After the approval from Hospital University Kebangsaan Malaysia's (HUKM) Ethics Committee, the EEG signals of 20 idiopathic PD patients (10 women and 10 men, average age 59.05 ± 5.64 years, range 45–65 years) were acquired. The *mean* duration of PD was 5.75 ± 3.52 years (range between 1 and 12 years). The Hoehn and Yahr (H-Y) severity stage was I–III: *Two* PD patients were in stage I, *eleven* were in stage II, and *seven* were in stage III. The Mini-Mental State Examination (MMSE) scores were within the normal limits (26.90 ± 1.51 [range 25–30]). Exclusion criteria included the presence of other neurological disorders (e.g., epilepsy) or psychiatric conditions (e.g., depression) and any other severe mental illness. All the PD patients took levodopa (L-dopa) drugs in order to reduce the heterogeneity in the medication.

Twenty age-matched normal subjects (11 women and 9 men, average age 58.10 ± 2.95 years) with no history or symptoms of neurological or mental illness served are recruited. The scores of the MMSE for normal subjects were 27.15 ± 1.63. Both PD patients and normal participants were self-reported as right-handed, confirmed by the Edinburgh Handedness Inventory (EHI), and without impairments of hearing. Approval for this study was sought from all participants by explaining the potential risk involved.

2.2 EEG recordings

The participants were seated comfortably in a quiet room with eyes-closed state to attain a state of relaxed wakefulness and were instructed prior to the study to avoid any body movements such as eye movement/blinking during the experiment. The EEG signals were recorded for 5 min in eyes-closed resting state on 14-channel (AF3, AF4, F3, F4, F5, F6, F7, F8, T7, T8, P7, P8, O1, and O2) wireless (2.4 GHz band) Emotiv EPOC neuroheadset at 128 Hz sampling rate.

3 Analysis of the EEG data

Figure 1 displays the block diagram of the proposed methodology. It consists of signal preprocessing, HOS features, ranking, integrated Parkinson's Disease Diagnosis Index (PDDI), and classification steps.

3.1 Preprocessing

The artifacts due to eye blinking were eliminated by thresholding technique during the preprocessing of EEG signals, by discarding the amplitudes of more than 80 μV. Then, the data were filtered using sixth-order Butterworth bandpass filter with lower and higher cutoff frequencies of 1 and 49 Hz in order to reduce the artifact components. The filtering is performed in forward and reverse, twice, to cancel the phase nonlinearity of the butterworth filter. In each channel, the artifact free signals were separated into EEG epochs of 2 s for further processing [24].

3.2 Feature extraction—higher-order spectra (HOS)

The HOS is a powerful tool used to study the nonlinear characteristics of the EEG signal [25]. It is a spectral representation of higher-order statistics. HOS has the ability to preserve the information due to deviations from Gaussianity and degrees of nonlinearities in the time series. As it is expected that the EEG signals have nonlinearities in the generating mechanism, the HOS analysis of PD patient's EEG might reveal additional non-Gaussian and nonlinear information due to its certain advantage [25]. In this work, the third-order spectra of the signal called the "bispectrum" were implemented. It is defined as: $B(f_1,f_2) = E[X(f_1)X(f_2)X^*(f_1+f_2)]$, where $B(f_1,f_2)$ is the bispectrum in the bifrequency (f_1,f_2), $X(f)$ is the discrete-time Fourier transform of the given signal, * denotes complex conjugate, and E[] denotes the statistical expectation operation over an ensemble of possible realizations of the signal [25]. The bispectrum is the most accessible of the HOS as it is the simplest to compute (computational complexity increases with increasing order) and its properties have been well explored [20, 25, 26]. The bispectrum display symmetry is evaluated in the principal domain region (Ω) as given in [25]. Bicoherence is the squared-magnitude of the normalized bispectrum [25]. A total of *thirteen* bispectrum features are extracted, namely bispectrum *mean* magnitude ($\mathrm{Bi}_{\mathrm{Mag}}$) [27], bispectral entropies (BiEnt_1 and BiEnt_2) [20], bispectrum phase entropy (BiPhEnt) [20], bispectrum moments [26] (sum of logarithmic amplitudes of bispectrum (H_1), sum of logarithmic amplitudes of diagonal elements in the bispectrum (H_2), first-order spectral moment of amplitudes of diagonal elements of the bispectrum (H_3), second-order spectral moment of amplitudes of diagonal elements of the bispectrum (H_4), first-order spectral moment of amplitudes of the principal domain region in the bispectrum (H_5), weighted center of bispectrum (WCB_{ix} and WCB_{iy}) [28] and absolute weighted center of bispectrum (WCB_{ix} and WCB_{iy}) [20]. These bispectrum features can capture the minute changes in the EEG signals to discriminate PD and healthy brain dynamics and can be used for automated diagnosis. To calculate these bispectrum features, epochs of 256 samples (2 s) with 50% overlap Hanning window and record of 256 NFFT points at 128 Hz sampling rate were used. The mathematical equation of thirteen extracted bispectrum features is subsequently given:

$$\mathrm{Bi}_{\mathrm{Mag}} = \frac{1}{N}\sum_{\Omega}|\mathrm{Bi}(f_1,f_2)| \tag{1}$$

$$\mathrm{BiEnt}_1 = -\sum_x q_x \log(q_x), \quad \text{where} \quad q_x = \frac{|\mathrm{Bi}(f_1,f_2)|}{\sum_{\Omega}|\mathrm{Bi}(f_1,f_2)|} \tag{2}$$

$$\mathrm{BiEnt}_2 = -\sum_x S_x \log(S_x), \quad \text{where} \quad S_x = \frac{|\mathrm{Bi}(f_1,f_2)|^2}{\sum_{\Omega}|\mathrm{Bi}(f_1,f_2)|^2} \tag{3}$$

$$\mathrm{BiPhEnt} = \sum_x \mathrm{ph}(\alpha_x)\log \mathrm{ph}(\alpha_x), \quad \text{where} \quad \mathrm{ph}(\alpha_x) = \frac{1}{N}\sum_{\Omega} l(\varphi(\mathrm{Bi}(f_1,f_2)) \in \alpha_x) \tag{4}$$

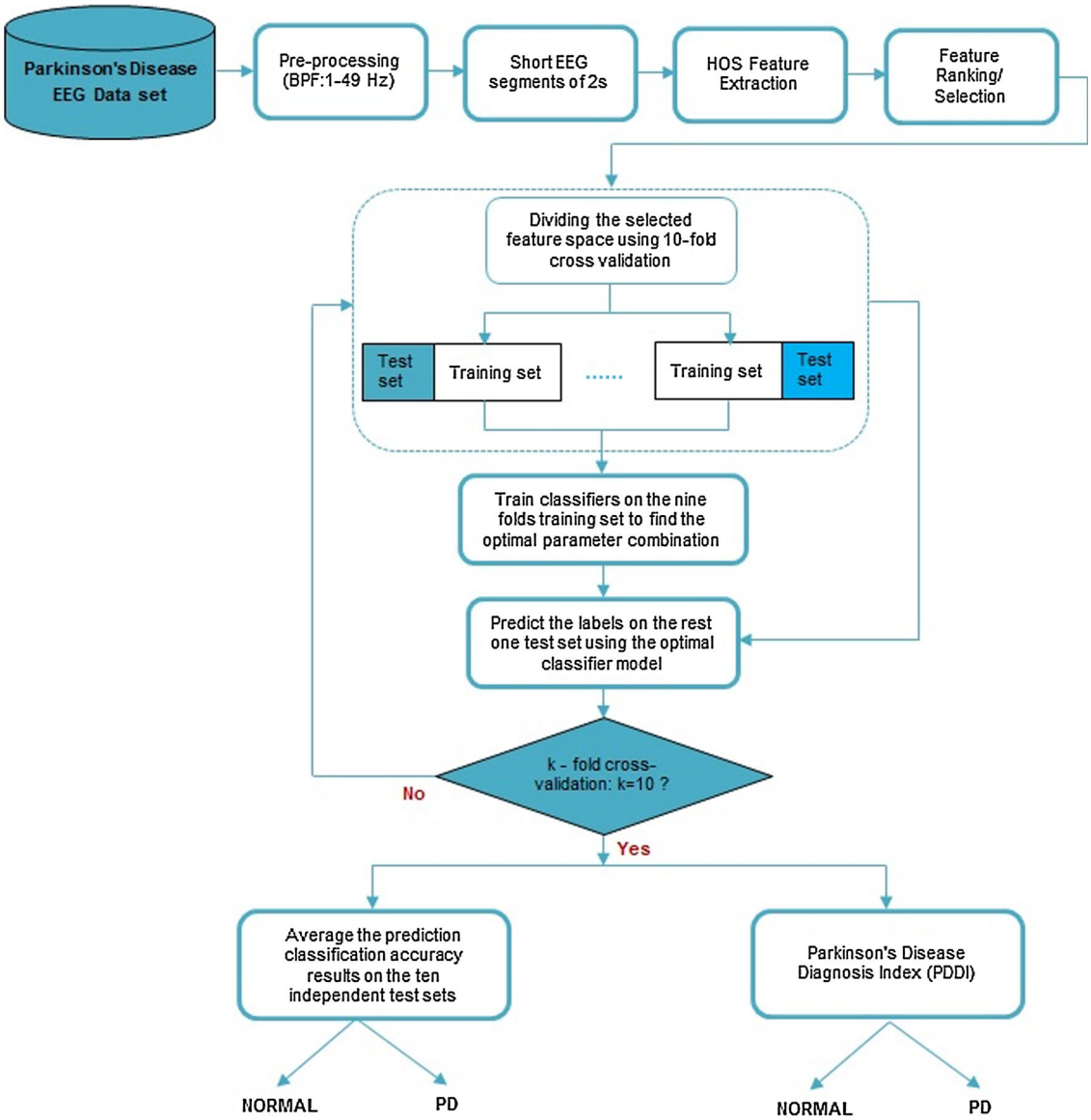

Fig. 1 System for automated identification of PD patients

$$\alpha_x = \{\phi | -\pi + 2\pi x/M \le \phi < -\pi + 2\pi(x+1)/M, x = 0, 1, 2, \ldots, M-1\}$$

$$H_1 = \sum\nolimits_{\Omega} \log(|\mathrm{Bi}(f_1, f_2)|) \tag{5}$$

$$H_2 = \sum\nolimits_{\Omega} \log(|\mathrm{Bi}(f_D, f_D)|) \tag{6}$$

$$H_3 = \sum_{m=1}^{N} m \log(|\mathrm{Bi}(f_D, f_D)|) \tag{7}$$

$$H_4 = \sum_{m=1}^{N} (m - H_3)^2 \log(|\mathrm{Bi}(f_D, f_D)|) \tag{8}$$

$$H_5 = \sum_{\Omega} \sqrt{i^2 + j^2} \log\left(\left|\mathrm{Bi}(f_i, f_j)\right|\right) \tag{9}$$

$$\mathrm{WCBi}x = \frac{\sum_{\Omega} i \mathrm{Bi}(i,j)}{\sum_{\Omega} \mathrm{Bi}(i,j)} \quad \text{and} \quad \mathrm{WCBi}y = \frac{\sum_{\Omega} j \mathrm{Bi}(i,j)}{\sum_{\Omega} \mathrm{Bi}(i,j)} \tag{10}$$

$$a\mathrm{WCBi}x = \frac{\sum_{\Omega} i |\mathrm{Bi}(i,j)|}{\sum_{\Omega} |\mathrm{Bi}(i,j)|} \quad \text{and} \quad a\mathrm{WCBi}y = \frac{\sum_{\Omega} j |\mathrm{Bi}(i,j)|}{\sum_{\Omega} |\mathrm{Bi}(i,j)|} \tag{11}$$

where $N =$ total number of points within Ω region, $\phi =$ phase angle of the bispectrm, $1(.) =$ function whose value will be 1 when the phase angle falls inside bin α_x, and i and j are bispectrum frequency bin index in the principal domain region, Ω.

3.3 Feature ranking/selection

The feature extraction step usually results in a more number of feature vectors, and many of these vectors may not contribute to differentiating the two classes. Hence, the most common idea is to apply ranking/selection algorithms on the extracted feature vectors to rank these feature vectors based on their discriminating ability. This as well reduces the complexity of classifiers without disturbing its performance. In this study, Student's *t* test is used for this purpose [29, 30]. For each feature vector, the *t* test yields 2 parameters: (i) the *p* value—it is used to determine the significance of the extracted features, a low *p* value indicates a high significance—and (ii) *t* value—it is used to rank the feature vectors, a high *t* value indicates better rank and feature. To evaluate the performance, the ranked features are added one by one to a particular classifier until the highest classification accuracy is reached.

3.4 Parkinson's Disease Diagnosis Index (PDDI)

It is time-consuming to develop an automated system as it involves extracting features, feature ranking, training, and testing. So, it is more convenient for the clinicians to use a single number that clearly separates the two classes. The concept of integrated index is first conceived by Ghista [31, 32] and further applied for the diagnosis of depression [33], epilepsy [34], sudden cardiac death [35], carotid plaque [36], thyroid [37], diabetes [38], and glaucoma [39]. Accordingly, we have proposed and formulated an integrated index called the Parkinson's Disease Diagnosis Index (PDDI), by combining most distinguishing feature vectors in such a way that the integrated index value is distinctly different for normal and PD patients.

The PDDI is developed using highly ranked *three* features (H_1, Ent_1, and H_2) from Table 1. We have developed a mathematical Eq. (12) by trial-and-error method in such a way that it clearly discriminates the two classes using a single number (PDDI). The mathematical formulation of this PD index is given by

$$\mathrm{PDDI} = \frac{\{(3.5 * \mathrm{ENT1}) + [0.5 * (H1/H2)]\}}{10} \tag{12}$$

3.5 Classification

In *decision tree* (DT) classifier, the input feature vectors were used to construct a tree [40]. This tree provides the rules to classify the two classes, and these rules were used to determine the test data class. The performance of this classifier depends on how well the tree was designed. The Gini index was used to measure the impurity at each node [41]. *Fuzzy K-nearest neighbor* (FKNN) classifier designates a class based on the major class among the KNN. Here, Euclidean distance was used in FKNN to allocate the fuzzy class membership before taking decisions. To determine how heavily the distance was weighted when calculating each neighbor's contribution to the membership value, fuzzy strength parameter (m) was used. In this study, the highest classification performance was obtained using $m = 1.24$ and $k = 3$. *K-nearest neighbor* (KNN) classifier calculates the minimum distance between testing and training data in terms of K-nearest neighbors [40]. The Euclidean distance and $k = 2$ were used for evaluating the separation. *Naive Bayes* (NB) classifier is a probabilistic classifier that works on Bayes theorem and on the assumption that the features are independent random variables [42]. *Probabilistic neural network* (PNN) is a multilayered feed forward network that uses the exponential activation function. In this study, the best performance was achieved using the smoothing parameter (σ) value 0.284. *Support vector machine* (SVM) classifier separates the training data into two classes in the feature space, by constructing a separating hyperplane [43]. The nonlinear signals which are not easily separable are converted to a higher-dimensional feature space using kernel functions. Polynomial kernel functions of order 2 and 3, radial basis function (RBF), and linear kernels were used in this work.

3.6 Performance measures

The tenfold cross-validation was used to evaluate the performance of the developed system, and the performance is evaluated using *five* different measures: *sensitivity*—true-positive value that quantifies the percentage of correctly classified PD patients among all PD feature vectors; *specificity*—true-negative value that quantifies the percentage of correctly classified normal subjects in all the healthy feature vectors; *accuracy*—% of correctly classified samples (both PD patients and normal subjects) in the total feature vectors; *precision*—% of correctly classified PD samples in all feature vectors recognized as PD; and ***F-score***—harmonic *mean* of the precision and sensitivity. Herein, 20 participants per group with 150 EEG segments of 2 s per electrode which resulted in a total feature vectors of 3000 × 14 (electrodes) were analyzed.

4 Results

Figure 2a, b displays the magnitude bispectrum plots of normal and PD EEG signals. It can be observed from these figures that the magnitude in the bifrequency plane is unique for each class. It is also clear that the bispectrum has most of its magnitude within −0.2 to +0.2 (the bifrequency range) in the normalized scale (i.e.,

Table 1 Mean ± SD of bispectrum feature vectors extracted from EEG signals for PD and normal subjects

HOS features	Normal subjects Mean ± SD	PD patients Mean ± SD	*p* value	*F* value	Ranking
H_1	4.88 E4 ± 3.06 E3	3.57 E4 ± 3.33 E3	2.00E−34	153.1306	1
Ent_1	0.943 ± 0.047	0.701 ± 0.165	1.59E−33	148.8198	2
H_2	1.18 E3 ± 1. E2	1.06 E2 ± 1.0 E2	2.63E−25	109.8145	3
H_3	2.0 E5 ± 1.7 E4	1.9 E5 ± 1.4 E4	7.55E−24	102.9533	4
aWcobx	113.193 ± 26.281	103.050 ± 30.215	2.27E−23	100.7048	5
H_4	6.6 E5 ± 2.1 E4	6.1 E5 ± 1.5 E4	1.68E−15	64.02431	6
Ent_2	0.546 ± 0.139	0.510 ± 0.190	8.84E−10	37.7809	7
PEnt	3.571 ± 0.066	3.579 ± 0.014	9.21E−05	15.3292	8
Wcobx	86.254 ± 77.287	79.300 ± 81.235	0.019159	5.492283	9
aWcoby	113.193 ± 26.281	35.970 ± 10.625	0.147529	2.84423	10
H_5	1.1 E6 ± 2.8 E5	1.0 E6 ± 1.8 E5	0.105925	2.615394	11
Wcoby	40.629 ± 45.177	38.096 ± 38.245	0.110908	2.542575	12
M_{avg}	2.6 E3 ± 1.9 E2	1.8 E3 ± 3.2 E2	0.122087	2.391545	13

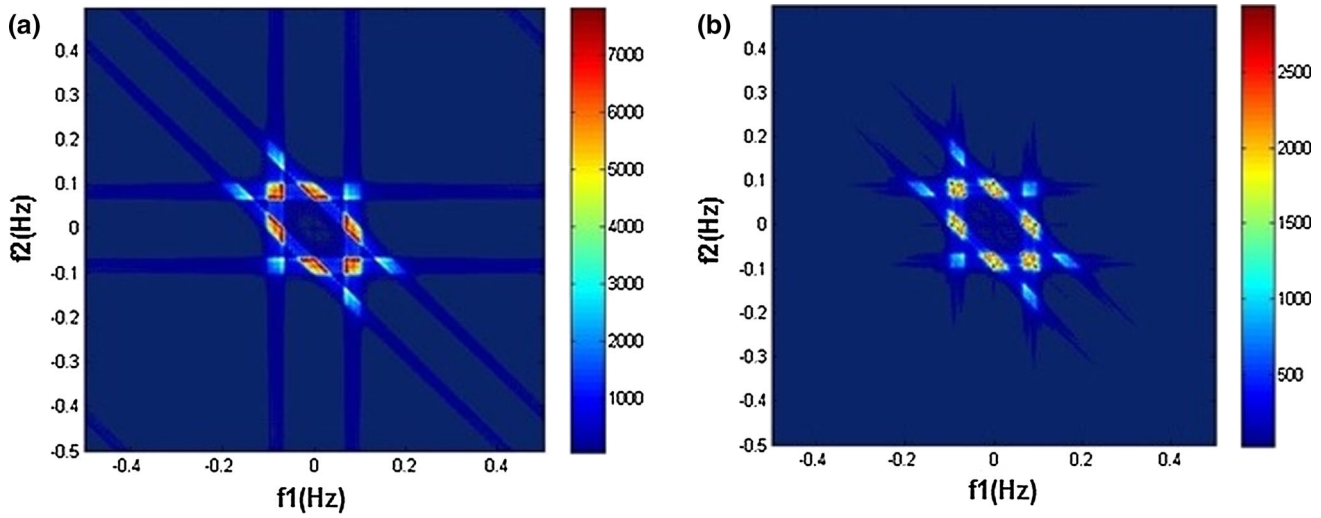

Fig. 2 Bispectrum plots of (**a**) PD patient EEG (**b**) normal subject EEG

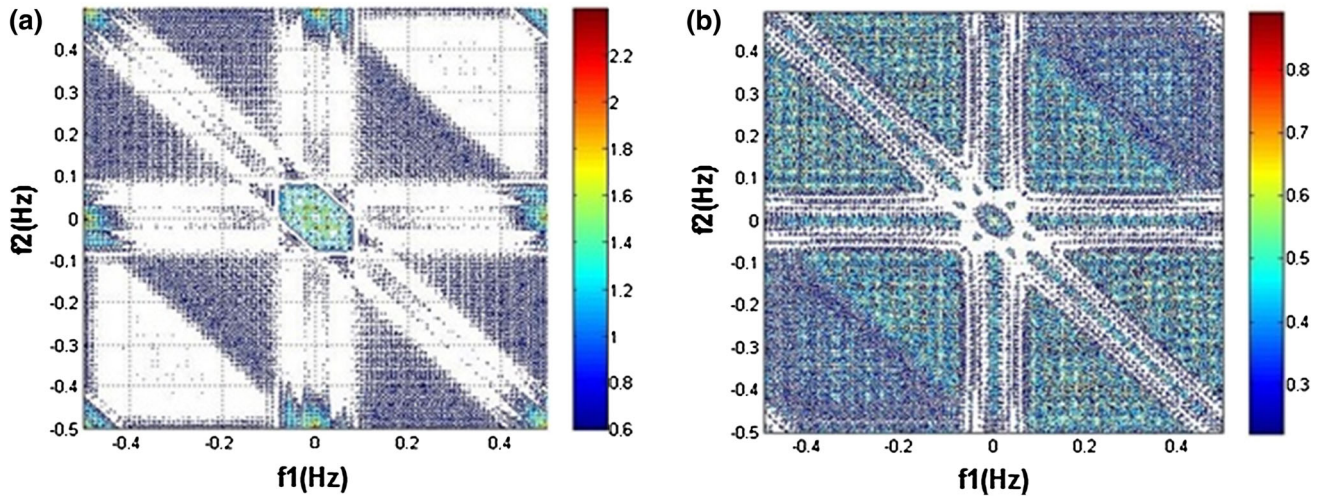

Fig. 3 Bicoherence plots of (**a**) PD patient EEG (**b**) normal subject EEG

Table 2 Performance measures of various classifiers using PD patients and normal subjects EEG signals based on different combinations of HOS features

Classifier	No. of HOS features	Performance measures (mean ± SD)				
		Sensitivity (%)	Specificity (%)	Precision (%)	*F*-score	Accuracy (%)
DT	6	94.17 ± 2.03	85.63 ± 1.05	94.17 ± 1.03	0.90 ± 0.06	90.65 ± 2.76
FKNN ($m = 1.24$, $k = 3$)	4	95.00 ± 1.83	90.21 ± 2.61	95.00 ± 1.83	0.93 ± 0.05	93.01 ± 2.89
KNN ($k = 2$)	5	96.67 ± 1.30	82.50 ± 2.72	96.67 ± 0.88	0.90 ± 0.05	90.83 ± 3.97
NB	7	94.17 ± 1.62	80.83 ± 3.82	94.17 ± 1.62	0.88 ± 0.06	88.72 ± 2.94
PNN ($\sigma = 0.284$)	5	96.21 ± 0.83	90.61 ± 1.24	96.67 ± 1.83	0.94 ± 0.05	93.98 ± 1.39
SVM linear	4	95.48 ± 1.83	87.58 ± 3.82	95.00 ± 1.83	0.92 ± 0.05	91.98 ± 2.95
SVM poly 2	4	96.74 ± 1.83	95.36 ± 1.83	96.67 ± 1.83	0.96 ± 0.04	96.07 ± 1.22
SVM poly 3	4	95.83 ± 1.39	95.83 ± 1.39	95.83 ± 1.39	0.95 ± 0.04	95.12 ± 1.28
SVM-RBF	3	100.00 ± 0.00	99.25 ± 0.53	99.38 ± 0.47	0.98 ± 0.05	99.62 ± 0.57

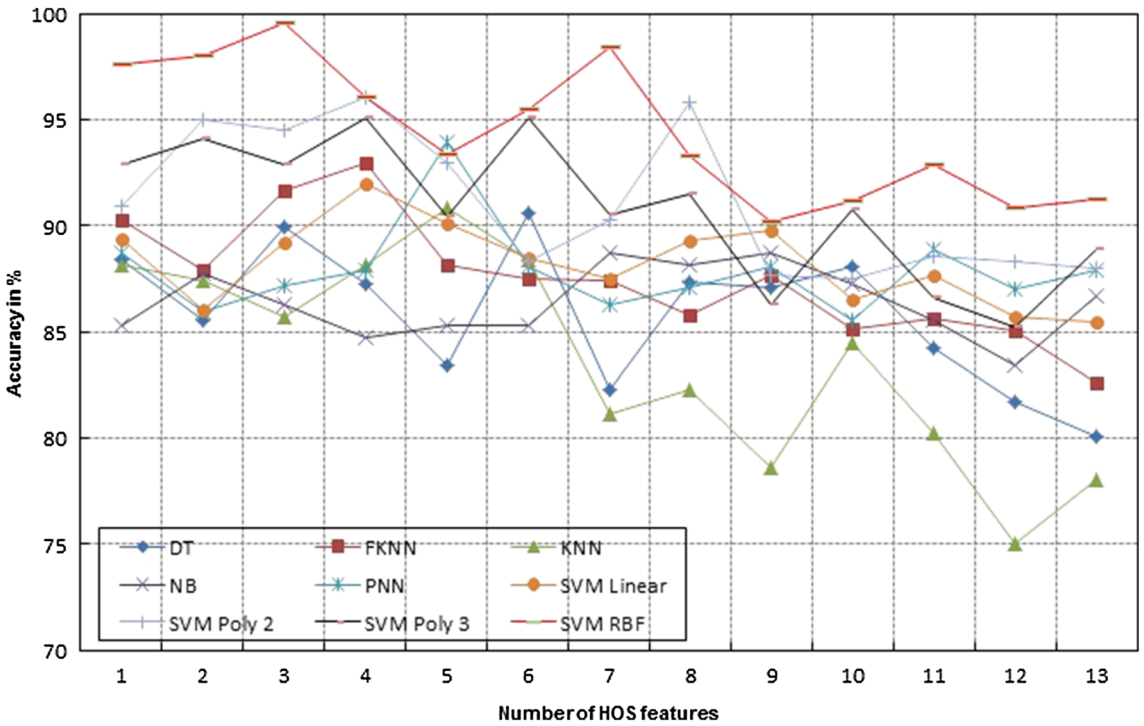

Fig. 4 Plot of number of features versus average accuracy for the various classifiers used

$-0.5 \leq f_1, f_2 \leq +0.5$). Figure 3a and b displays the bicoherence plots of normal subjects and PD patient's EEG signals. These plots also indicate that the magnitudes are randomly distributed at various frequencies throughout the plot. In normal subject, the spread in the bifrequency plane is more compared to PD patient. This may be due to the fact that, as the person becomes diseased, the EEG signal becomes less chaotic.

In this work, a total of *thirteen* bispectrum features from EEG segments of 2 s were extracted. Table 1 presents the mean ± standard deviation (SD) of various HOS feature vectors. The feature vectors were ranked with respect to their *t* values. It can be noted from Table 1 that all the extracted bispectrum features are decreased for PD class. This observation suggests that the EEG signal became less complex for PD patients due to dysfunction in the neural circuits. The *p* value and *t* value indicated that almost all the features were clinically significant. Table 2 shows the results of classification using different classifiers. It can be noted that the SVM classifier using RBF kernel function (SVM-RBF) achieved an optimum *mean* accuracy of 99.62 ± 0.57%, sensitivity of 100.00 ± 0.00%, and specificity of 99.25 ± 0.53%. This performance was obtained using only *three* ranked bispectrum features, namely H_1, Ent_1, and H_2. The EEG signals are nonlinear in nature; hence, the nonlinear kernel functions, such as RBF, perform well. Figure 4

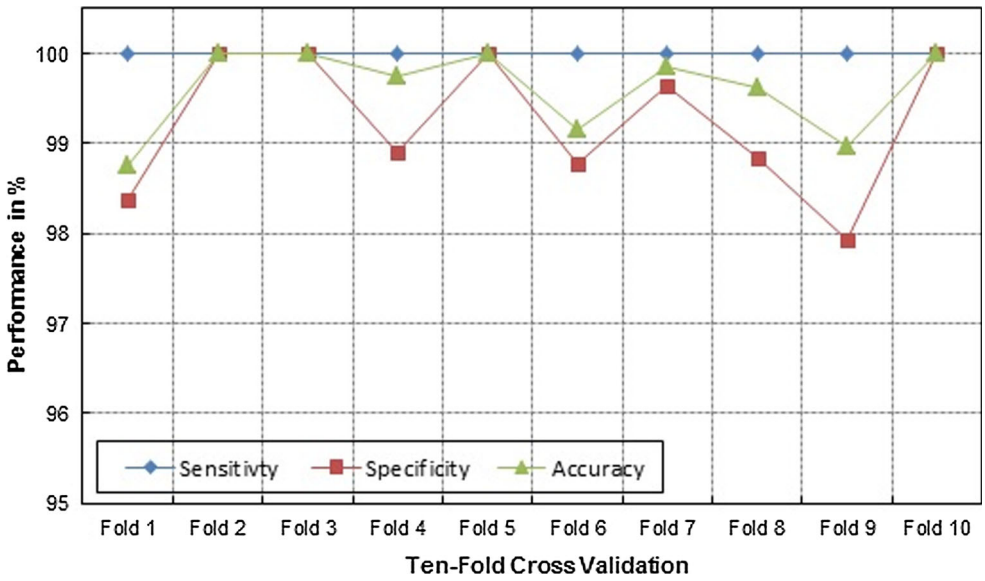

Fig. 5 Plot of average accuracy (%), sensitivity (%), specificity (%) vs number of folds (tenfold cross-validation for SVM-RBF classifier)

Table 3 Range of PDDI for normal and PD classes

	Normal	PD	*p* value
PDDI	14.6035 ± 0.55922	10.60263 ± 0.681133	<0.0001

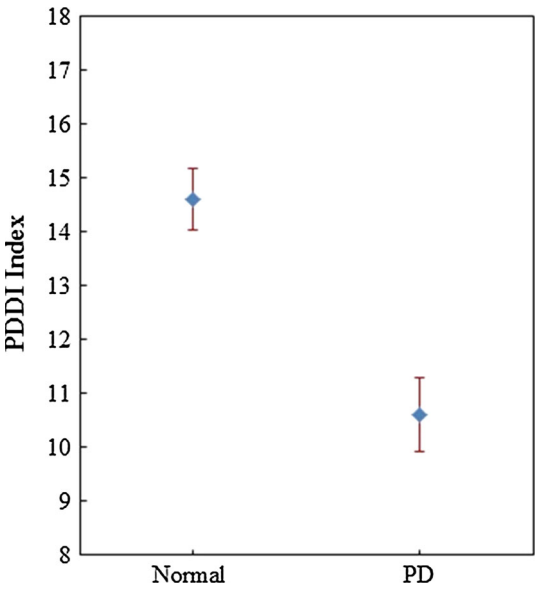

Fig. 6 Plot of PDDI for normal subjects and PD patients

displays the plot of accuracy vs number of features for various classifiers used. It clearly shows that the classifiers yielded better classification accuracy for the top *three* ranked features, beyond which there is a decrease in accuracy level. Figure 5 shows the plot of *mean* accuracy (%), sensitivity (%), and specificity (%) vs number of folds (tenfold cross-validation) using SVM-RBF classifier.

Table 3 shows the range of PDDI for normal and PD patients. Figure 6 shows the plot of PDDI for two classes. The table and figure enable to understand that there is a clear separation between the two classes, and hence, we can separate them using a single number.

5 Discussion

In this study, a nonlinear method for an automated diagnosis of PD was proposed using HOS bispectrum features extracted from EEG signals. The novelty of this work is the formulation of PDDI and also proposed unique HOS plots for normal and PD classes. Table 4 lists the classification accuracies of works conducted in the diagnosis of PD. It can be understood from the table that most of the studies have used dysphonia-based features (group of vocal symptoms) to identify the PD. However, it is well known that PD is predominantly a motor disorder mainly caused due to the loss of dopamine-producing neurons in the basal ganglia. Also, non-motor impairment involving cognitive dysfunction in PD patients has often been noted [44]. Cognitive status is mainly associated with neurophysiological signals (e.g., EEGs). Subsequently, an understanding of the neuronal activity is important to compare voice impairment symptoms, for the advancement of both targeted therapeutic strategies and prognostic perspectives. It has been revealed that the analysis of EEG signals can help to show the disturbed subcortico-cortical mechanisms in PD patients or dementia [9, 10, 12]. Thus, this study was performed to develop an automated detection system using EEG signals for PD patients. In addition, it can be

Table 4 Summary of studies conducted on automated detection of PD and normal classes

Authors	Signal used	Features	Method	Classifier	Accuracy (%)
Little et al. (2009) [45]	Voice (UCI machine learning database)	HNR, RPDE, DFA, and PPE	Preselection filter + exhaustive search	SVM	91.4 (bootstrap with 50 replicates)
Shahbaba and Neal (2009) [46]	Voice (UCI machine learning database)	Raw dysphonia features	Dirichlet process mixtures	–	87.70 (fivefold CV)
Das (2010) [47]	Voice (UCI machine learning database)	Raw data samples	ANN	Back-propagation learning algorithm	92.9 (hold-out)
Sakar and Kursun (2010) [48]	Voice (UCI machine learning database)	Mutual information	Mutual information-based feature selection	SVM	92.75 (bootstrap with 50 replicates)
Psorakis et al. (2010) [49]	Voice (UCI machine learning database)	mRVM	Improved mRVMs	SVM	89.47 (tenfold CV)
Guo et al. (2010) [50]	Voice (UCI machine learning database)	Decision boundaries	GP-EM	–	93.10 (tenfold CV)
Ozcift and Gulten (2011) [51]	Voice (UCI machine learning database)	Raw data samples	CFS	RF	87.1 (tenfold CV)
Li et al. (2011) [52]	Voice (UCI machine learning database)	Raw data samples	Fuzzy-based nonlinear transformation	SVM	93.47 (hold-out)
Luukka (2011) [53]	Voice (UCI machine learning database)	Raw data samples	Fuzzy entropy measures	Similarity classifier	85.03 (hold-out)
Spadoto et al. (2011) [54]	Voice (UCI machine learning database)	Raw data samples	PSO + OPF; Harmony search + OPF; gravitational search	–	73.53 (hold-out); 84.01 (hold-out); 84.01 (hold-out)
Astrom and Koker (2011) [55]	Voice (UCI machine learning database)	Raw data samples	Parallel NN	–	91.20 (hold-out)
Ozcift (2012) [56]	Voice (UCI machine learning database)	HNR, RPDE, DFA, and PPE	RF ensemble of IBk (a *k*-nearest neighbor variant) algorithm	–	97.00
Polat (2012) [57]	Voice (UCI machine learning database)	Raw data weighting	FCMFW	KNN	97.93 (50–50% training–testing)
Tsanas et al. (2012) [58]	Voice (UCI machine learning database)	Raw dysphonia features	Feature selection: LASSO, mRMR, RELIF, LLBFS	SVM	Almost 90.00% (tenfold CV)
Daliri [5]	Gait (from underneath of subject's feet)	Moment, mean and variance frequency	STFT + FDR	SVM	91.20 (50–50% training–testing)
Chen et al. (2013) [59]	Voice (UCI machine learning database)	Dataset normalization (0, 1)	PCA	FKNN	96.07 (average tenfold CV)
Zuo et al. [7]	Voice (UCI machine learning database)	Dataset normalization (−1, 1)	PSO	FKNN	97.47 (tenfold CV)
Ma et al. [4]	Voice (UCI machine learning database)	Cluster centers mean	SCFW	KELM	99.49 (tenfold CV)
This work	EEG segments of 2 s	Bispectrum features	HOS + feature ranking	SVM	99.62 (average tenfold CV) *PDDI

* Parkinson's Disease Diagnosis Index

seen from the table that researchers have achieved accuracy between 70 and 99.5% using various methods.

Table 4 shows that the HOS-based method, proposed by the authors of this paper, has given the superior performance compared to all available modalities. Moreover, the best distinguishing features were selected and combined into an integrated index as shown in Eq. (1), to optimally separate the two classes.

The main salient features of proposed automated diagnosis system are as follows:

1. The method yielded an optimum *mean* accuracy: 99.62%, sensitivity: 100%, and specificity: 99.25% using bispectrum features.
2. The method proposed PDDI using short segments (2 s) of EEG signals. This can be used by the clinicians for the diagnosis of the PD using one numeric value.
3. The method was implemented using MATLAB software and can be installed in hospitals. This could help to reduce the workload of neurologists and to help the process of accurate diagnosis of the PD patients.

4. The developed technique is fully automatic, noninvasive, and robust.
5. Extracted HOS-based bispectrum features are more robust to noise, and the method can be extended to the diagnosis of other neurological disorders such as cerebral palsy etc.

The sensitivity of 100% and specificity of 99.25% using only *three* highly ranked HOS features were obtained. However, the proposed method needs to be tested with a larger database of PD patients belonging to diverse ethnic groups. This necessitates a large storage space to extract the features and run the classification algorithms. The study was conducted using Intel i7-2410 M processors @2.30 GHZ with 8.0 GB RAM. The entire experiment was carried out using MATLAB (version 8.1.0.604, R2014a). The average time required to train and test the system was 6.484 s.

The limitation of this study is that only 20 PD patients were participated and the data were only obtained from Malaysian race. In order to get a reliable index, huge dataset with data from diverse ethnic groups is needed.

6 Conclusion

This study presented an application of HOS features extracted from EEG signals for diagnosis of PD patients. The findings demonstrate that the proposed technique is able to discriminate PD from normal EEG signals using a single number (PDDI) clearly and also with an average sensitivity, specificity, and accuracy of 100, 99.25, and 99.62%, respectively, using SVM classifier. Thereby, our developed EEG-based automated system can be used as a promising alternative tool in the diagnosis of PD. The proposed index provides a distinct non-overlapping ranges for normal and PD classes. It can help the neurologists in faster and more accurate diagnosis of PD during their screening. The proposed technique can be extended in order to classify the severity levels of PD.

Acknowledgements The authors would like to thank all the individuals who supported and participated in this study.

Compliance with ethical standards

Conflict of interest The authors declare that there is no conflict of interest.

References

1. Han CX, Wang J, Yi GS, Che YQ (2013) Investigation of EEG abnormalities in the early stage of Parkinson's disease. Cogn Neurodyn 7:351–359
2. Valls-Sole J, Valldeoriola F (2002) Neurophysiological correlate of clinical signs in Parkinson's disease. Clin Neurophysiol 113(6):792–805
3. Hossen A, Muthuraman M, Raethjen J, Deuschl G, Heute U (2010) Discrimination of Parkinsonian tremor from essential tremor by implementation of a wavelet-based soft-decision technique on EMG and accelerometer signals. Biomed Signal Process Control 5:181–188
4. Ma C, Ouyang J, Chen HL, Zhao XH (2014) An efficient diagnosis system for Parkinson's disease using kernel-based extreme learning machine with subtractive clustering features weighting approach. Comput Math Methods Med 2014:1–14. doi:10.1155/2014/985789
5. Daliri MR (2013) Chi square distance kernel of the gaits for the diagnosis of Parkinson's disease. Biomed Signal Process Control 8(1):66–70
6. Chen HL, Huang CC, Yub XG, Xu X, Sun X, Wang G, Wang SJ (2013) An efficient diagnosis system for detection of Parkinson's disease using fuzzy k-nearest neighbor approach. Expert Syst with Appl 40(1):263–271
7. Zuo WL, Wang ZY, Liu T, Chen HL (2013) Effective detection of Parkinson's disease using an adaptive fuzzy K-nearest neighbor approach. Biomed Signal Process Control 8(4):364–373
8. Neufeld MY, Inzelberg R, Korczyn AD (1988) EEG in demented and non-demented parkinsonian patients. Acta Neurol Scand 78(1):1–5
9. Neufeld MY, Blumen S, Aitkin I, Paramet Y, Korczyn AD (1994) EEG frequency analysis in demented and non-demented parkinsonian patients. Dementia (Basel, Switzerland) 5(1):23–28
10. Pezard L, Jech R, Ruzicka E (2001) Investigation of non-linear properties of multichannel EEG in the early stages of Parkinson's disease. Clin Neurophysiol 112:38–45
11. Soikkeli R, Partanen J, Soininen H, Pääkkönen A, Riekkinen PS (1991) Slowing of EEG in Parkinson's disease. Electroencephalogr Clin Neurophysiol 79(3):159–165
12. Tanaka H, Koenig T, Pascual-Marqui RD, Hirata K, Kochi K, Lehmann D (2000) Event-related potential and EEG measures in Parkinson's disease without and with dementia. Dement Geriatr Cogn Disord 11(1):39–45
13. Muller V, Lutzenberger W, Pulvermüller F, Mohr B, Birbaumer N (2001) Investigation of brain dynamics in Parkinson's disease by methods derived from nonlinear dynamics. Exp Brain Res 137(1):103–110
14. Lima CAM, Coelho ALV, Chagas S (2009) Automatic EEG signal classification for epilepsy diagnosis with Relevance Vector Machines. Expert Syst Appl 36(6):10054–10059
15. Leuchter AF, Cook IA, Gilmer WS, Marangell LB, Burgoyne KS, Howland RH, Trivedi MH, Zisook S, Jain R, Fava M, Iosifescu D, Greenwald S (2009) Effectiveness of a quantitative electroencephalographic biomarker for predicting differential response or remission with escitalopram and bupropion in major depressive disorder. Psychiatry Res 169(2):132–138
16. Gandal MJ, Edgar JC, Klook K, Siegel SJ (2012) Gamma synchrony: towards a translational biomarker for the treatment-resistant symptoms of schizophrenia. Neuropharmacology 62(3):1504–1518
17. Hampel H, Frank R, Broich K, Teipel SJ, Katz RG, Hardy J, Herholz K, Bokde AL, Jessen F, Hoessler YC, Sanhai WR, Zetterberg H, Woodcock J, Blennow K (2010) Biomarkers for Alzheimer's disease: academic, industry and regulatory perspectives. Nat Rev Drug Discov 9(7):560–574
18. Al-Hazimi A, Al-Ama N, Syiamic A, Qosti R, Abdel-Galil K (2002) Time-domain analysis of heart rate variability in diabetic patients with and without autonomic neuropathy. Ann Saudi Med 22(5–6):400–403
19. Stam CJ (2005) Nonlinear dynamical analysis of EEG and MEG: review of an emerging field. Clin Neurophysiol 116:2266–2301

20. Chua KC, Chandran V, Acharya UR, Lim CM (2009) Analysis of epileptic EEG signals using higher order spectra. J Med Eng Technol 33(1):42–50
21. Acharya UR, Chua EC, Chua KC, Min LC, Tamura T (2010) Analysis and automatic identification of sleep stages using higher order spectra. J Neural Syst 20(6):509–521
22. Martis RJ, Acharya UR, Mandana KM, Ray AK, Chakraborty C (2013) Cardiac decision making using higher order spectra. Biomed Signal Process Control 8:193–203
23. Yuvaraj R, Murugappan M, Norlinah MI, Sundaraj K, Omar MI, Khairiyah M, Palaniappan R (2014) Optimal set of EEG features for emotional state classification and trajectory visualization in Parkinson's disease. Int J Psychophysiol 94(3):482–495
24. Aydin S, Arica N, Ergul E, Tan O (2014) Classification of obsessive compulsive disorder by EEG complexity and hemispheric dependency measurements. Int J Neural Syst 25(3):1–16
25. Chua KC, Chandran V, Acharya UR, Lim CM (2010) Application of higher order statistics/spectra in biomedical signals–a review. J Med Eng Phy 32:679–689
26. Zhou SM, Gan JQ, Sepulveda F (2008) Classifying mental tasks based on features of higher order statistics from EEG signals in brain computer interface. Info Sci 178:1629–1640
27. Du X, Dua S, Acharya RU, Chua CK (2012) Classification of epilepsy using high-order spectra features and principal component analysis. J Med Syst 36(3):1731–1743
28. Zhang JW, Zheng CX, Xie A (2000) Bispectrum analysis of focal ischemic cerebral EEG signal using third-order recursion method. IEEE Trans Biomed Eng 47(3):352–359
29. Kobayashi H, Mark BL, Turin W (2011) Probability, random processes and statistical analysis: applications to communications, signal processing queuing theory and mathematical finance. Cambridge University Press, Cambridge
30. Acharya UR, Fujita H, Sudarshan VK, Bhat S, Koh JEW (2015) Application of entropies for automated diagnosis of epilepsy using EEG signals: a review. Knowledge-Based Syst 88:85–96
31. Ghista DN (2004) Physiological systems numbers in medical diagnosis and hospital cost effective operation. J Mech Med Biol 4:401–418
32. Ghista DN (2009) Non-dimensional physiological indices for medical assessment. J Mech Med Biol 9:643–669
33. Acharya UR, Sudarshan VK, Adeli H, Santhosh J, Koh JE, Puthankatti SD, Adeli A (2015) A novel depression diagnosis index using nonlinear features in EEG signals. Eur Neurol 74(1–4):79–83
34. Sharma R, Pachori RB, Acharya UR (2015) An integrated index for the identification of focal electroencephalogram signals using discrete wavelet transform and entropy measures. Entropy 17(8):5218–5240
35. Acharya UR, Fujita H, Vidya KS, Vinitha S, Lim WJE, Ghista DN, Tan RS (2015) An integrated index for detection of sudden cardiac death using discrete wavelet transform and nonlinear features. Knowledge-Based Syst 83:149–158
36. Acharya UR, Faust O, Alvin APC, Sree SV, Molinari F, Saba L, Nicolaides A, Suri JS (2012) Symptomatic vs. asymptomatic plaque classification in carotid ultrasound. J Med Syst 36(3):1861–1871
37. Acharya UR, Faust O, Sree SV, Molinari F, Garberoglio R, Suri JS (2011) Cost-effective and non-invasive automated benign and malignant thyroid lesion classification in 3D contrast-enhanced ultrasound using combination of wavelets and textures: a class of Thyro Scan TM algorithms. Tech Cancer Res Treat 10(4):371–380
38. Acharya UR, Vidya KS, Ghista DN, Lim WJE, Molinari F, Meena S (2015) Computer aided diagnosis of diabetic subjects by heart rate variability signals using discrete wavelet transform method. Knowledge-Based Syst 81:56–64
39. Mookiah MRK, Acharya UR, Lim CM, Petznick A, Suri JS (2012) Data mining technique for automated diagnosis of glaucoma using higher order spectra and wavelet energy features. Knowledge-Based Syst 33:73–82
40. Larose DT (2004) Discovering knowledge in data: an introduction to data mining. Willey Interscience, New Jersey, pp 90–106 **(Chapter 5)**
41. Jerritta S, Murugappan M, Wan K, Yaacob S (2013) Classification of emotional states from electrocardiogram signals: a nonlinear approach based on hurst. Biomed Eng Online 12(1):44–62
42. Han J, Kamber M (2006) Data mining: concepts and techniques, 2nd edn. Morgan Kaufmann, Burlington
43. Vapnik VN (1999) An overview of statistical learning theory. IEEE Trans Neural Netw 10(5):988–999
44. Cooper JA, Sagar HJ, Jordan N, Harvey NS, Sullivan EV (1991) Cognitive impairment in early, untreated Parkinson's disease and its relationship to motor disability. Brain 114(Pt 5):2095–2122
45. Little MA, McSharry PE, Hunter EJ, Spielman J, Ramig LO (2009) Suitability of dysphonia measurements for telemonitoring of Parkinson's disease. IEEE Trans Biomed Eng 56(4):1015–1022
46. Shahbaba B, Neal R (2009) Nonlinear models using Dirichlet process mixtures. J Machine Learning Res 10:1829–1850
47. Das R (2010) A comparison of multiple classification methods for diagnosis of Parkinson disease. Expert Syst Appl 37(2):1568–1572
48. Sakar CO, Kursun O (2010) Telediagnosis of Parkinson's disease using measurements of dysphonia. J Med Syst 34(4):591–599
49. Psorakis I, Damoulas T, Girolami MA (2010) Multiclass relevance vector machines: sparsity and accuracy. IEEE Trans Neural Netw 21(10):1588–1598
50. Guo PF, Bhattacharya P, Kharma N (2010) Advances in detecting Parkinson's disease. In: Medical biometrics, Lecture Notes in Computer Science, Berlin, Germany 6165:306–314
51. Ozcift A, Gulten A (2011) Classifier ensemble construction with rotation forest to improve medical diagnosis performance of machine learning algorithms. Comput Methods Progr Biomed 104(3):443–451
52. Li DC, Liu CW, Hu SC (2011) A fuzzy-based data transformation for feature extraction to increase classification performance with small medical data sets. Artif Intell Med 52(1):45–52
53. Luukka P (2011) Feature selection using fuzzy entropy measures with similarity classifier. Expert Syst Appl 38(4):4600–4607
54. Spadoto AA, Guido RC, Carnevali FL, Pagnin AF, Falcao AX, Papa JP (2011) Improving Parkinson's disease identification through evolutionary-based feature selection. In: Proceedings of the annual international conference of the ieee engineering in medicine and biology society (EMBC-11), Boston, Mass, USA, pp. 7857–7860
55. Astrom F, Koker R (2011) A parallel neural network approach to prediction of Parkinson's Disease. Expert Syst Appl 38(10):12470–12474
56. Ozcift A (2012) SVM feature selection based rotation forest ensemble classifiers to improve computer-aided diagnosis of Parkinson disease. J Med Syst 36:2141–2147
57. Polat K (2012) Classification of Parkinson's disease using feature weighting method on the basis of fuzzy C-means clustering. Int J Syst Sci 43(4):597–609
58. Tsanas A, Little MA, McSharry PE, Spielman J, Ramig LO (2012) Novel speech signal processing algorithms for high-accuracy classification of Parkinson's disease. IEEE Trans Biomed Eng 59(5):1264–1271
59. Chen HL, Huang CC, Yu XG, Xuc X, Sund X, Wang G, Wang SJ (2013) An efficient diagnosis system for detection of Parkinson's disease using fuzzy *k*-nearest neighbor approach. Expert Syst Appl 40(1):263–271